

E.U

DE 98/03366  
PRIORITY  
DOCUMENT

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



REC'D	20 JAN 1999
WIPO	PCT

### Bescheinigung

Die Siemens Aktiengesellschaft in München/Deutschland hat eine  
Patentanmeldung unter der Bezeichnung

"Verfahren und Vorrichtung zur Spracherkennung"

am 21. November 1997 beim Deutschen Patent- und Markenamt einge-  
reicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe  
der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig  
die Symbole G 10 L und G 06 F der Internationalen Patentklassi-  
fikation erhalten.

München, den 30. November 1998  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

*Wehner*

Patenzzeichen: 197 51 739.0

Wehner

This Page Blank (uspto)

~~Ursprungsfassung 19751739.0~~

1

Beschreibung**Verfahren und Vorrichtung zur Spracherkennung**

- 5 Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Spracherkennung.

Ein Spracherkennungssystem ist aus [1] bekannt. Dort finden sich auch eine grundlegende Einführung der an dem  
10 Spracherkennungssystem beteiligten Komponenten sowie wichtiger, bei der Spracherkennung üblicher Techniken.

Bei einem bekannten Spracherkennungssystem ist eine Genauigkeit, also ein Maß für eine Qualität der Erkennung,  
15 vorgegeben. Der Benutzer muß nun mit diesem System auskommen, auch wenn für seine Anwendung eine verminderte Genauigkeit ausreichen, dafür er aber eine höhere Bearbeitungsgeschwindigkeit erzielen würde.

20 Das Prinzip des Pruning eines Suchraums ist aus [2] (siehe Kapitel 3.3.3, Seite 40) bekannt. Dabei handelt es sich um ein "Beschneiden" des Suchraums, also eine Methode zur Reduzierung einer Anzahl von Suchpfaden des Suchraums, wobei diejenigen Suchpfade abgeschnitten werden, die wenig aussichtsreich sind. Zuerst wird dazu ein Suchpfad mit minimalen Kosten (optimaler Suchpfad) ermittelt. Daraufhin werden alle Suchpfade (Äste des Suchbaums) weggeschnitten, deren Kosten oberhalb des Minimums zuzüglich einer addierten vorgegebenen Bewertungsgröße, die als Pruning-Schwelle  
30 bezeichnet wird, liegen. Für eine detaillierte Erklärung des Prunings: [2], Seite 40ff., insbesondere Bild 16 auf Seite 41.

Bei Verwendung der Pruning-Schwelle ist nicht bekannt, wieviel Suchpfade in dem Suchbaum übrig bleiben. Will man die  
35 Anzahl dieser übrigbleibenden Suchpfade auf einem vorgegebenen Niveau halten, wird die Pruning-Schwelle dynamisch angepaßt.

Ein Histogramm-Pruning ist aus [3] oder [4] bekannt. Hier werden eine vorgegebene Anzahl "bester" Suchpfade, also Suchpfade mit einer hohen Auftrittswahrscheinlichkeit, verwendet, indem Häufigkeiten der Suchpfade in Form eines Histogramms bewertet werden. Die Pruning-Schwelle wird dynamisch verändert.

10 Eine akustische Vorausschau im Suchbaum (Fachwort: Fast-Look-Ahead) ist aus [5] oder [6] bekannt.

Die bei der akustischen Vorausschau (auch schnelle Vorauswahl) verfolgte Idee beruht auf der Eigenschaft einer Sprache, daß sich alle Wörter aus einem beschränkten Inventar von Unterworteinheiten (z.B. Phonemen, Halbsilben) zusammensetzen. Für diese Unterworteinheiten wird nun "im Voraus" eine akustische Bewertung durchgeführt. Es werden nur diejenigen Kombination von Unterworteinheiten weiterverfolgt, deren akustische Bewertungen jeweils unterhalb einer vorgegebenen Schwelle liegen. Ein Gewinn im Bewertungsaufwand besteht darin, daß für eine geringe Anzahl von Unterworteinheiten ein Maß für die Übereinstimmung eines zu erkennenden Sprachsignals mit einer Zielgröße im voraus berechnet und als Grundlage für eine Entscheidung herangezogen wird, ob ein Großteil des Suchbaums nicht weiter berücksichtigt werden soll. Anschaulich gesprochen bedeutet dies, daß mehr Suchpfade im Suchbaum eingespart werden, als durch die Vorausberechnung hinzukommen. Ein derartiger Gewinn wird umso größer, je höher das Verhältnis von neuen Wortanfängen zu einer Anzahl von Unterworteinheiten wird. Dieses Verhältnis steigt mit der Anzahl der zu erkennenden Unterworteinheiten bzw. Wörter (Lexikongröße).

Ein Vorteil des Verfahrens der akustischen Vorausschau besteht in der Regularität der Algorithmen zur Berechnung der entsprechenden Maße. Da keine Verzweigungen durch Wortenden, Syntaxknoten, etc. im Suchraum auftreten, ist das Schema der Berechnung der Maße regulär. Gerade deshalb bietet sich ein

solches Verfahren auch für eine Implementierung in Hardware an.

Die Vorausberechnung der Maße (Fachwort: Fast-Match-Scores) wird dadurch möglich, daß die eigentliche Suche um eine feste Anzahl von Zeitfenstern hinter den aktuellen extrahierten Maßen des Sprachsignals hinterhereilt. Mit den aktuellen Maßen wird die Vorausberechnung der Maße weiterer Unterworteinheiten durchgeführt (siehe [5], Seite 65, Bild 33).

Auch in Sprachmodellen ist die Durchführung einer derartigen Vorausschau bekannt (siehe [6]).

Das Prinzip der Vorausschau im Sprachmodell (Fachwort: Language-Model-Look-Ahead) ist die Berücksichtigung der im Sprachmodell vorhandenen Wahrscheinlichkeiten in dem Suchprozeß so früh wie möglich, auch in dem assoziierten Pruning. Dies wird erreicht durch eine Faktorisierung der Wahrscheinlichkeiten im Sprachmodell. Eine detaillierte Beschreibung mit einer formalen Notation ist in [6] enthalten.

Schließlich ist z.B. aus [7] eine Schwelle zur Auswahl berechnender Distanzparameter bekannt. Derartige Auswahlverfahren sind generell mehrstufig angelegt. Zuerst wird eine grobe Berechnung mit einem Teil der Distanzen durchgeführt. Im nächsten Schritt werden dann diejenigen Distanzen bestimmt, die bezüglich eines Abstandsmaßes nahe an der besten Distanz des ersten Berechnungsschrittes liegen. Dieses Abstandsmaß kann über eine Schwelle variiert werden, wodurch der Berechnungsaufwand für die Bestimmung der Distanzparameter variiert wird.

Die Aufgabe der Erfindung besteht darin, ein Verfahren und eine Vorrichtung zur Einstellung der Genauigkeit des Spracherkennungssystems zu schaffen.

Diese Aufgabe wird gemäß den Merkmalen der unabhängigen Patentansprüche gelöst.

Es wird ein Verfahren zur Einstellung einer Genauigkeit eines  
5 Spracherkennungssystems angegeben, bei dem durch eine  
vorgebbare Eingangsgröße die Genauigkeit bestimmt wird.  
Anhand dieser Eingangsgröße werden Werte für Systemparameter  
des Spracherkennungssystems, vorzugsweise mittels eines  
Rechners, ermittelt. Anhand dieser Werte wird das  
10 Spracherkennungssystem eingestellt. Dies geschieht  
vorzugsweise automatisch durch den Rechner.

Somit ist es ein Vorteil der Erfindung, die Genauigkeit des  
Spracherkennungssystems adaptierbar und zugleich für den  
15 Laien einstellbar zu machen. Je nach Anwendung bzw. je nach  
Rechenleistung, die für das Spracherkennungssystem auf dem  
Rechner bereitsteht, können unterschiedliche Anforderungen an  
die Qualität des Spracherkennungssystems einfach durch  
Adaption der Eingangsgröße vorgenommen werden.

20 Eine Weiterbildung besteht darin, daß gemäß einer  
Abbildungsvorschrift aus der Eingangsgröße die Werte für die  
Systemparameter des Spracherkennungssystems ermittelt werden.  
Dabei kann diese Abbildungsvorschrift anhand einer Tabelle  
25 umgesetzt sein.

Es ist also möglich, durch Einstellung der Eingangsgröße  
automatisch die damit verknüpften Werte der Systemparameter  
des Spracherkennungssystems zu ermitteln und automatisch dem  
30 Spracherkennungssystem zugänglich zu machen. Das Ablegen der  
Werte in einer Tabelle hat den Vorteil, daß eine individuelle  
Anpassung verschiedener Werte der Eingangsgröße zu jeweils  
verschiedenen Werten der Systemparameter durchgeführt werden  
kann.

35 Eine andere Weiterbildung besteht darin, die Einstellung  
während des Betriebs des Spracherkennungssystems

durchzuführen. Dabei ergibt sich vorteilhaft, daß die Anpassung des Spracherkennungssystems individuell nach den jeweiligen Anforderungen während des Betriebs des Spracherkennungssystems angepaßt werden kann.

5

Eine zusätzliche Weiterbildung der Erfindung besteht darin, daß mindestens einer der folgenden Systemparameter anhand der Eingangsgröße bestimmt wird:

- a) Pruning-Schwelle;
- 10 b) Histogramm-Pruning;
- c) akustische Vorausschau;
- d) Vorausschau im Sprachmodell;
- e) Schwelle zur Auswahl zu berechnender Distanzparameter.

15 Zur Bedeutung und Funktion dieser Systemparameter wird auf die Einleitung verwiesen.

Die angeführten Systemparameter stellen eine Auswahl von Möglichkeiten dar. Es sind andere Systemparameter  
20 vorstellbar, die, abhängig von dem jeweiligen Spracherkennungssystem, von den oben genannten verschieden sein können.

Auch ist es eine Weiterbildung der Erfindung, daß die Systemparameter hinsichtlich ihres Einflusses auf eine Zielgröße hin gewichtet werden. Dabei kann die Zielgröße beispielsweise eine Genauigkeit des Spracherkennungssystems oder eine Geschwindigkeit des Spracherkennungssystems (also die Geschwindigkeit zur Durchführung des  
30 Spracherkennungsprozesses) sein. Im Hinblick auf die jeweilige Zielgröße können die Systemparameter anteilig gleich oder entsprechend einer vorgegebenen Gewichtungstabelle unterschiedlich gewichtet werden.

35 Im Rahmen einer anderen Weiterbildung wird die Eingangsgröße anhand eines Einstellelements bestimmt.

Vorzugsweise weist das Einstellelement einen eindimensionalen Freiheitsgrad mit zwei Begrenzungen auf, wobei die erste Begrenzung als maximale Genauigkeit des Spracherkennungssystems und die zweite Begrenzung als maximale Geschwindigkeit des Spracherkennungssystems umgesetzt werden.

Auch ist es eine Weiterbildung des Verfahrens, daß das Einstellelement auf einem Rechner als ein Schieberegler dargestellt wird und anhand einer Tastatur, eines Touch-Pads oder einer Maus bedient wird.

Auch kann das Einstellelement ein Drehregler, ein Schieberegler oder Potentiometer sein.

In einer zusätzlichen Weiterbildung wird das Einstellelement über Sprache angesteuert, die von einem Spracherkenner, insbesondere dem Spracherkennungssystem, ausgewertet wird. Die Eingangsgröße kann mittels Spracheingabe bestimmt werden.

Das Verfahren wird auch weitergebildet, indem eine vollständig automatisierte Bestimmung der Eingangsgröße in den folgenden Schritten durchgeführt wird:

Eine Leistungsfähigkeit des Rechners, auf dem das Spracherkennungssystem ablaufen soll, wird anhand eines Programms zur Leistungsermittlung bestimmt und als ein Leistungsindex abgespeichert. Unter Berücksichtigung des Leistungsindex werden die Systemparameter des Spracherkennungssystems automatisch eingestellt und dadurch wird eine leistungsfähige Spracherkennung unter z.B. Echtzeitbedingung gewährleistet.

Ein Programm zur Bestimmung der Leistungsfähigkeit des Rechners kann eine vorgegebene Schleife sein, die eine bestimmte Anzahl Iterationen durchläuft, wobei die Zeit für die Iterationen gemessen wird. Anhand einer Tabelle kann abhängig von der gemessenen Zeit der Leistungsindex bestimmt werden. Auch gibt es kommerziell oder als Freeware verfügbare



Programme, die die Leistungsfähigkeit des Rechners ermitteln und als eine Bewertungsgröße der Leistungsfähigkeit einen Leistungsindex ausgeben.

- 5 Auch wird eine Vorrichtung zur Spracherkennung angegeben, die ein Spracherkennungssystem aufweist und mit einem Mittel zur Einstellung einer Genauigkeit des Spracherkennungssystems ausgeführt ist, welches Mittel Systemparameter des Spracherkennungssystem aus einer Eingangsgröße umsetzt, also  
10 die Einstellung des Spracherkennungssystems und seiner zahlreichen Systemparameter anhand der Eingangsgröße vornimmt.

- Dabei ist es vorteilhaft, daß eine solche Einstellung der  
15 Systemparameter anhand der Eingangsgröße während des Betriebs des Spracherkennungssystems erfolgen kann. So wird für den Benutzer eine einfache Anpassung der Vielzahl von Systemparametern möglich.

- 20 Die Eingangsgröße ist in einer Weiterbildung automatisch bestimmbar. Dazu wird anhand einer Einrichtung zur Leistungsmessung des Rechners, auf dem das Spracherkennungssystem abläuft, ein sog. Leistungsindex ermittelt und zur Einstellung der Genauigkeit des Spracherkennungssystems benutzt.

- Eine andere Weiterbildung sieht vor, daß die Eingangsgröße durch ein Einstellelement vorgebar ist. Dazu sind eine Vielzahl möglicher Einstellelemente (Potentiometer, virtuelle  
30 Steuereinheiten auf dem Rechner, etc.) denkbar, deren Einstellung direkt die Genauigkeit des Spracherkennungssystems bestimmt.

- Vorteilhaft können Schritte des erfindungsgemäßen Verfahrens  
35 auf der angegebenen Vorrichtung durchgeführt werden.

Weiterbildungen der Erfindung ergeben sich auch aus den abhängigen Ansprüchen.

5 Anhand der folgenden Figuren werden Ausführungsbeispiele der Erfindung näher dargestellt.

Es zeigen

- 10 Fig.1 eine Systemarchitektur für ein  
Spracherkennungssystem;
- Fig.2 ein Blockdiagramm mit Schritten eines Verfahrens zur  
Einstellung der Genauigkeit eines  
15 Spracherkennungssystems;
- Fig.3 ein Blockdiagramm, das eine Verknüpfung einer  
Eingangsgröße mit mindestens einem Systemparameter  
über eine Abbildungsvorschrift darstellt;
- 20 Fig.4 ein Diagramm, das verschiedene mögliche  
Systemparameter des Spracherkennungssystems  
darstellt;
- Fig.5 eine Skizze, die den Einfluß der Systemparameter auf  
25 eine Zielgröße darstellt;
- Fig.6 ein Diagramm, das verschiedene Möglichkeiten zur  
Einstellung der Eingangsgröße darstellt;
- 30 Fig.7 ein Spracherkennungssystem mit einem Mittel zur  
Einstellung der Genauigkeit.

35 In Fig.1 ist allgemein eine Systemarchitektur für eine  
Spracherkennung (Spracherkennungssystem) dargestellt.

Voraussetzung für die Erkennung natürlich gesprochener  
Sprache ist ein geeigneter Formalismus zur

Wissensrepräsentation. Ein vollständiges  
Spracherkennungssystem umfaßt mehrere Verarbeitungsebenen.  
Dies sind insbesondere Akustik-Phonetik, Intonation, Syntax,  
Semantik und Pragmatik. In Fig.1 werden die  
5 Verarbeitungsebenen bei der Erkennung aufgezeigt.

Das natürliche Sprachsignal 101 gelangt in das  
Spracherkennungssystem. Dort wird in einer Komponente 102  
eine Merkmalsextraktion durchgeführt. Nach der  
10 Merkmalsextraktion werden anhand bekannter akustisch-  
phonetischer Einheiten 103 Sprachlaute erkannt (siehe Block  
104). Dabei handelt es sich um die Berechnung akustischer  
Distanzparameter. Nach der Sprachlauterkennung 104 erfolgt  
die lexikalische Decodierung (Worterkennung) in einem Block  
15 106 mit Hilfe des Aussprachemodells bzw. Wortlexikons 105 und  
daran anschließend eine Syntaxanalyse 108 mit Hilfe des  
Sprachmodells, das die Grammatik umfaßt, 107. Die  
Worterkennung 106 und die Syntaxanalyse 108 stellen die Suche  
nach einer Entsprechung für das Sprachsignal dar. Schließlich  
20 wird in einem Block 110 eine semantische Nachbearbeitung  
durchgeführt, wobei Kontextwissen und Pragmatik 109  
berücksichtigt werden und schließlich die vom  
Spracherkennungssystem erkannte Sprache 111 folgt.

In Fig.2 ist ein Blockdiagramm dargestellt, das Schritte  
eines Verfahrens zur Einstellung der Genauigkeit eines  
Spracherkennungssystems zeigt.

In einem Schritt 201 wird durch eine vorgebbare Eingangsgröße  
30 die Genauigkeit des Spracherkennungssystems bestimmt.  
Daraufhin wird in einem Schritt 202 anhand dieser  
Eingangsgröße für Systemparameter des Spracherkennungssystems  
Werte ermittelt. Schließlich wird in einem Schritt 203 das  
Spracherkennungssystem anhand der ermittelten Werte  
35 eingestellt.

**Fig.3** zeigt ein Blockdiagramm, das eine Verknüpfung einer Eingangsgröße über eine Abbildungsvorschrift mit mindestens einem Systemparameter darstellt.

Die erwähnte Eingangsgröße 301 wird anhand einer Abbildungsvorschrift 302 auf die Systemparameter SP 303 des Spracherkennungssystems abgebildet. Dabei wird vorzugsweise eine Eingangsgröße 301 mehreren Systemparametern über die Abbildungsvorschrift zugeordnet. Durch diese Abbildungsvorschrift 302 wird durch Vorgabe einer Eingangsgröße das Spracherkennungssystem angepaßt, also werden mehrere Systemparameter SP durch Veränderung einer Eingangsgröße 301 beeinflußt. Die Abbildungsvorschrift 302 hat vorzugsweise die Form einer Tabelle, in der eine Spalte mögliche Eingangsgrößen 301 enthält, und in einer Zeile dieser Spalte der jeweiligen Eingangsgröße mehrere Werte für Systemparameter SP 303 zugeordnet werden. Die Abbildungsvorschrift 302 besteht im Suchen nach dem der Eingangsgröße 301 zugeordneten Eintrag (Zeile) in der Tabelle und in der Übergabe der gefundenen Werte für Systemparameter SP 303 an das Spracherkennungssystem.

**Fig.4** zeigt ein Diagramm, das verschiedene mögliche Systemparameter des Spracherkennungssystems darstellt. Die Systemparameter SP des Spracherkennungssystems, dargestellt in einem Block 401, umfassen mindestens einen der folgenden Parameter:

- a) Pruning-Schwelle 402;
- b) Histogramm-Pruning 403;
- c) akustische Vorausschau 404;
- d) Vorausschau im Sprachmodell 405;
- e) Schwelle für Distanzparameter 406.

Insgesamt sind weitere Systemparameter des Spracherkennungssystems zur Einstellung über die Eingangsgröße 301 denkbar, angedeutet durch den Block 407.

Fig.5 zeigt eine Skizze, die den Einfluß der Systemparameter auf eine Zielgröße darstellt.

Die Systemparameter SP (siehe Block 501) nehmen Einfluß auf eine Zielgröße ZG (siehe Block 502). Wie oben beschrieben, gibt es mehrere Systemparameter SP1, SP2, usw., dargestellt in einem Block 503. Dabei nimmt jeder einzelne Systemparameter SP<sub>i</sub> (i=1,2,...) mittels eines für den jeweiligen Systemparameter SP<sub>i</sub> vorgesehenen Gewichts G<sub>i</sub> (dargestellt in einem Block 504) Einfluß auf die Zielgröße ZG. Durch die Gewichtung der Systemparameter SP ist es möglich, je nach Systemparameter SP einen unterschiedlichen Einfluß auf die Zielgröße ZG zu nehmen.

In Fig.6 ist ein Diagramm dargestellt, das verschiedene Möglichkeiten zur Einstellung der Eingangsgröße zeigt. Die Einstellung der Eingangsgröße, dargestellt in dem Block 601, erfolgt anhand von Bedienkomponenten des Rechners R (siehe Block 602), anhand von Regelungskomponenten 603 oder anhand des Rechners selbst (siehe Block 604). Der Rechner R umfaßt dazu mindestens ein Mittel zur Einstellung der Eingangsgröße, wie eine Tastatur 605, eine Maus 606, ein Touch-Pad 607 oder Spracheingabe 608 über das Spracherkennungssystem. Mögliche Komponenten zur Regelung, dargestellt in dem Block 603 sind ein Drehregler 609, ein Schieberegler 610 oder ein sonstiger Regler 611, vorzugsweise ein Potentiometer. Zusätzlich wird automatisch anhand eines Programms, das auf dem Rechner läuft, die Rechenleistung des Rechners ermittelt und die Eingangsgröße zur Einstellung des Spracherkennungssystems entsprechend bestimmt. Dadurch wird gewährleistet, daß ein automatisch eingestellter Rechner eine seiner Rechenleistung entsprechende Qualität bei der Spracherkennung gewährleistet. Automatisch wird ein Kompromiß gefunden zwischen hoher Qualität bei der Spracherkennung, unter Einbuße von Rechenleistung bzw. Echtzeiterkennung der Sprache, und schneller Spracherkennung mit entsprechend wenig Bedarf an

Rechenleistung allerdings deutlicher Qualitätseinbuße bei der Spracherkennung.

Fig.7 zeigt eine Vorrichtung aus einem Spracherkennungssystem 701 und einem Mittel zur Einstellung der Genauigkeit des Spracherkennungssystems 702.

Anhand des Mittels zur Einstellung der Genauigkeit 702 sind durch eine Eingangsgröße Systemparameter des Spracherkennungssystems bestimmt. Dazu wird die Eingangsgröße, vorzugsweise anhand einer vorgegebenen Tabelle, einer Vielzahl von Systemparametern (Pruning-Schwelle, Histogramm-Pruning, akustische Vorausschau, Vorausschau im Sprachmodell, Schwelle zur Auswahl zu berechnender Distanzparameter, etc.) zugeordnet.

Die Eingangsgröße ist wahlweise anhand eines Einstellelements 703 oder einer Einrichtung zur Leistungsbestimmung des Rechners 704 bestimmbar. Hierbei sei auch auf Fig.6 und die dort angegebenen Möglichkeiten zur Einstellung der Eingangsgröße verwiesen.

Im Rahmen dieses Dokuments wurden folgende Veröffentlichungen zitiert:

- 5 [1] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 2, Seiten 13 bis 26.
- [2] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 3.3.3, Seiten 40 bis 43.
- 15 [3] Volker Steinbiss, Bach-Hiep Tran, Hermann Ney: "Improvements in Beam Search. Proc. Intl. Conf. Speech and Language Processing, Yokohama 1994, Seiten 2143 bis 2146.
- 20 [4] M. Niemöller, A. Hauenstein, E. Marschall, P. Witschel, U. Harke: "A PC-based Real-Time Large Vocabulary Continuous Speech Recognizer for German", Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; München 1997.
- 25 [5] A. Hauenstein: "Optimierung von Algorithmen und Entwurf eines Prozessors für die automatische Spracherkennung", Lehrstuhl für Integrierte Schaltungen, Technische Universität München, Dissertation, 19.07.1993, Kapitel 3.5.1, Seiten 65 bis 69.
- 30 [6] S. Ortmanns, A. Eiden, H. Ney, N. Coenen: "Look-Ahead Techniques for Fast Beam Search", Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; München 1997, Seiten 1783 bis 1786.
- [7] E. Bocchieri: "Vector Quantization for the Efficient Computation of Continuous Density Likelihoods", Proc.

IEEE Intl. Conf. on Acoustics, Speech and Signal  
Processing; 1993, Seiten II-692 bis II-695.



Patentansprüche

1. Verfahren zur Spracherkennung,  
bei dem gesprochene Sprache anhand eines  
Spracherkennungssystems erkannt wird, wobei
  - a) anhand einer vorgebbaren Eingangsgröße eine Genauigkeit bestimmt wird;
  - b) anhand der Eingangsgröße Werte für Systemparameter des Spracherkennungssystems ermittelt werden;
  - c) anhand der Werte die Genauigkeit des Spracherkennungssystem eingestellt wird.
2. Verfahren nach Anspruch 1,  
bei dem die Werte für die Systemparameter des Spracherkennungssystems bestimmt werden, indem gemäß einer Abbildungsvorschrift aus der Eingangsgröße die Werte ermittelt werden.
3. Verfahren nach Anspruch 2,  
bei dem die Abbildungsvorschrift anhand einer Tabelle umgesetzt wird.
4. Verfahren nach einem der vorhergehenden Ansprüche,  
bei dem die Einstellung während des Betriebs des Spracherkennungssystems durchgeführt wird.
5. Verfahren nach einem der vorhergehenden Ansprüche,  
bei dem das Spracherkennungssystem mindestens einen der folgenden Systemparameter umfaßt:
  - a) Pruning-Schwelle;
  - b) Histogramm-Pruning;
  - c) akustische Vorausschau;
  - d) Vorausschau im Sprachmodell;
  - e) Schwelle zur Auswahl zu berechnender Distanzparameter.

6. Verfahren nach Anspruch 5,  
bei dem mindestens einer der Systemparameter anhand der  
Eingangsgröße bestimmt wird.
- 5 7. Verfahren nach Anspruch 6,  
bei dem die Systemparameter gewichtet werden hinsichtlich  
ihres Einflusses auf jeweils eine Zielgröße.
8. Verfahren nach Anspruch 7,  
10 bei dem eine Zielgröße mindestens eine der folgenden  
Größen ist:  
a) Genauigkeit des Spracherkennungssystems;  
b) Geschwindigkeit des Spracherkennungssystems.
- 15 9. Verfahren nach Anspruch 7 oder 8,  
bei dem die Systemparameter gleich gewichtet werden.
10. Verfahren nach Anspruch 7 oder 8,  
bei dem die Systemparameter entsprechend einer  
20 vorgegebenen Gewichtungstabelle gewichtet werden.
11. Verfahren nach einem der vorhergehenden Ansprüche,  
bei dem die Eingangsgröße anhand eines Einstellelements  
bestimmt wird.  
25
12. Verfahren nach Anspruch 11,  
bei dem das Einstellelement einen eindimensionalen  
Freiheitsgrad mit zwei Begrenzungen aufweist, dessen  
erste Begrenzung als maximale Genauigkeit des  
30 Spracherkennungssystems und dessen zweite Begrenzung als  
maximale Geschwindigkeit des Spracherkennungssystems  
umgesetzt werden.
13. Verfahren nach Anspruch 11 oder 12,  
35 bei dem das Einstellelement auf einem Rechner als ein  
Schieberegler dargestellt und anhand einer Tastatur,  
eines Touch-Pads oder einer Maus bedient wird.

14. Verfahren nach Anspruch 11 oder 12,  
bei dem das Einstellelement ein Drehregler, Schieberegler  
oder ein Potentiometer ist.

5

15. Verfahren nach Anspruch 11 oder 12,  
bei dem das Einstellelement über Sprache angesteuert  
wird, die von einem Spracherkenner, insbesondere dem  
Spracherkennungssystem, ausgewertet wird.

10

16. Verfahren nach einem der Ansprüche 1 bis 10,  
a) bei dem das Spracherkennungssystem auf einem Rechner  
abläuft;  
b) bei dem durch ein Programm zu einer  
15 Leistungsermittlung des Rechners ein Leistungsindex  
des Rechners bestimmt wird;  
c) bei dem die Eingangsgröße für das  
Spracherkennungssystem automatisch anhand des  
Leistungsindex bestimmt wird, wobei dabei eine  
20 Rechenleistung des Rechners automatisch auf die  
Genauigkeit des Spracherkennungssystems eingestellt  
wird.

20

17. Vorrichtung zur Spracherkennung,  
a) bei der ein Spracherkennungssystem vorgesehen ist,  
b) bei der ein Mittel zur Einstellung einer Genauigkeit  
des Spracherkennungssystems vorgesehen ist, das derart  
eingerrichtet ist, daß Systemparameter des  
Spracherkennungssystems einstellbar sind, wobei die  
30 Systemparameter anhand einer Eingangsgröße ermittelbar  
sind.

30

18. Vorrichtung nach dem vorherigen Anspruch,  
mit einer Einrichtung zur Leistungsmessung, die derart  
35 eingerichtet ist, daß die Eingangsgröße automatisch  
bestimmbar ist.

35

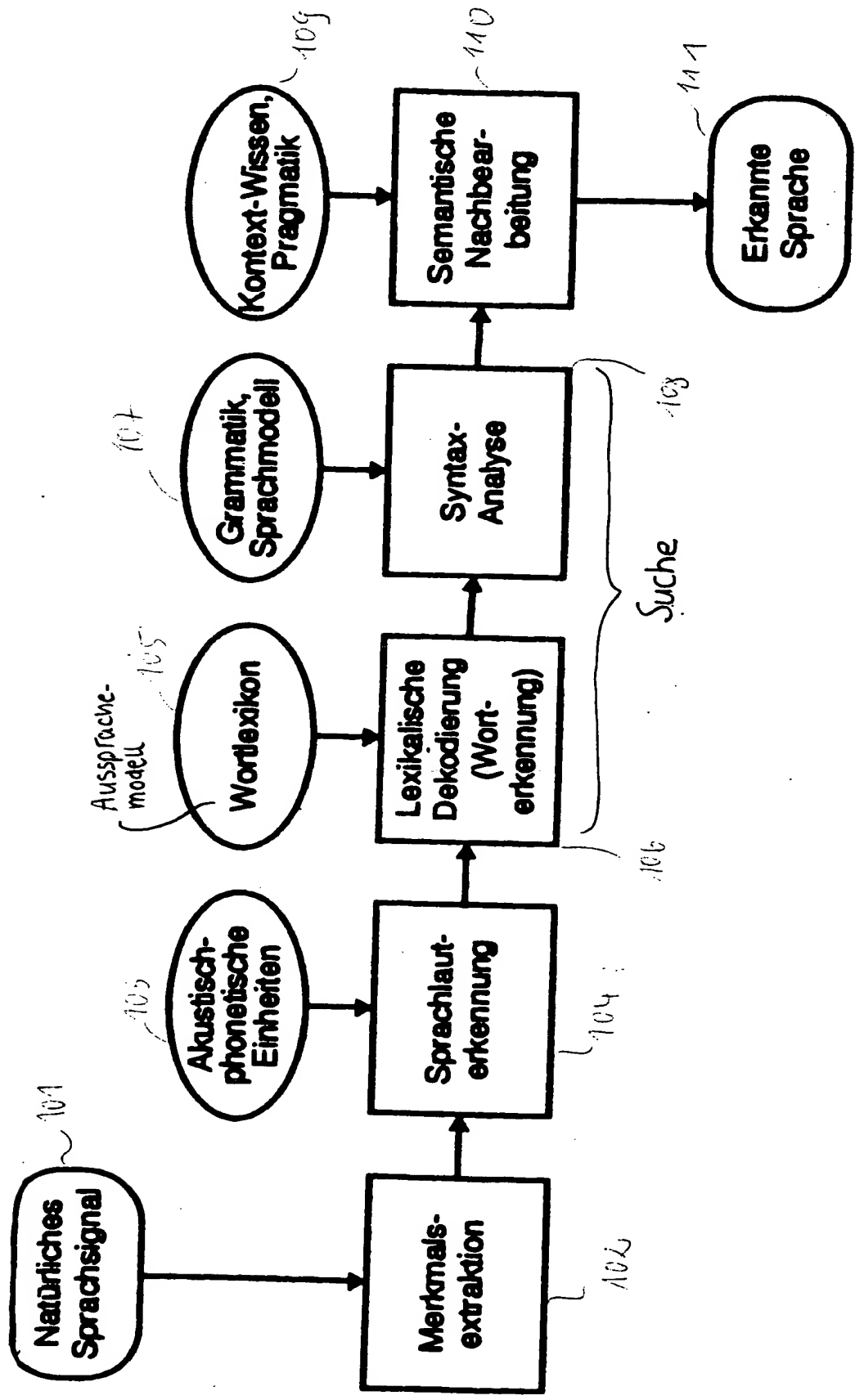
19. Vorrichtung nach Anspruch 17,  
mit einem Einstellelement, das derart eingerichtet ist,  
daß die Eingangsgröße durch das Einstellelement vorgebbar  
ist.

Zusammenfassung

## Verfahren und Vorrichtung zur Spracherkennung

- 5 Es werden ein Verfahren und eine Vorrichtung angegeben, die bei einer Spracherkennung eine Einstellung einer Genauigkeit eines Spracherkennungssystems ermöglichen. Dazu werden anhand einer Eingangsgröße über eine Abbildungsvorschrift Systemparameter des Spracherkennungssystems ermittelt und das
- 10 Spracherkennungssystem entsprechend der auf diese Art bestimmten Werte eingestellt. Optional kann eine Adaption eines Spracherkennungssystems während des Betriebs erfolgen, um eine anwendungsabhängige Anpassung in einem Bereich zwischen maximaler Qualität bei der Spracherkennung und
- 15 möglichst hoher Geschwindigkeit bei der Durchführung der Spracherkennung zu gewährleisten. Auch kann automatisch der Rechner eingestellt werden, indem zuvor anhand eines dafür vorgesehenen Programmes ein Leistungsindex des Rechners bestimmt wurde, der ein Maß für die Eingangsgröße darstellt
- 20 und somit einen adäquaten Betrieb des Spracherkennungssystems auf diesem Rechner gewährleistet.

Fig. 1



System-Architektur für die Spracherkennung

Bestimmung der  
Eingangsgröße

201



Erstellung der  
Werte der Synthesparameter  
anhand der Eingangsgröße

202



Erstellung des  
Sprachkennungssystems

203

FIG 3

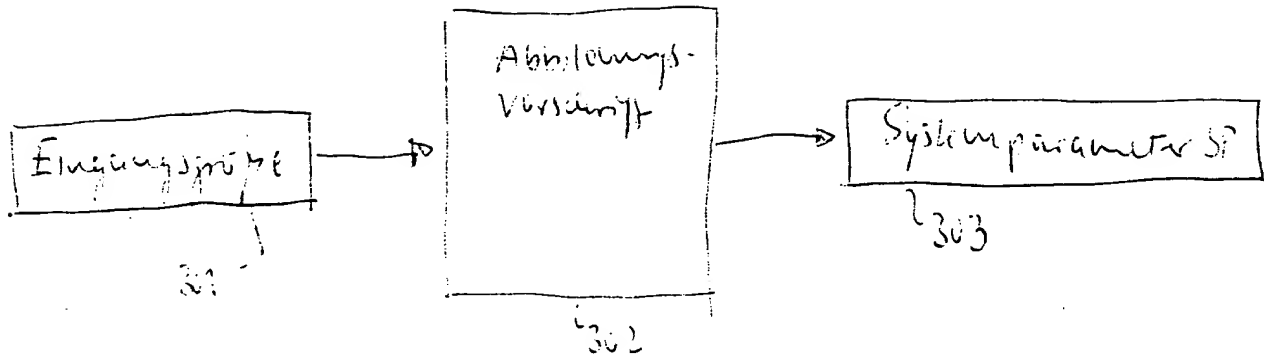


FIG 4

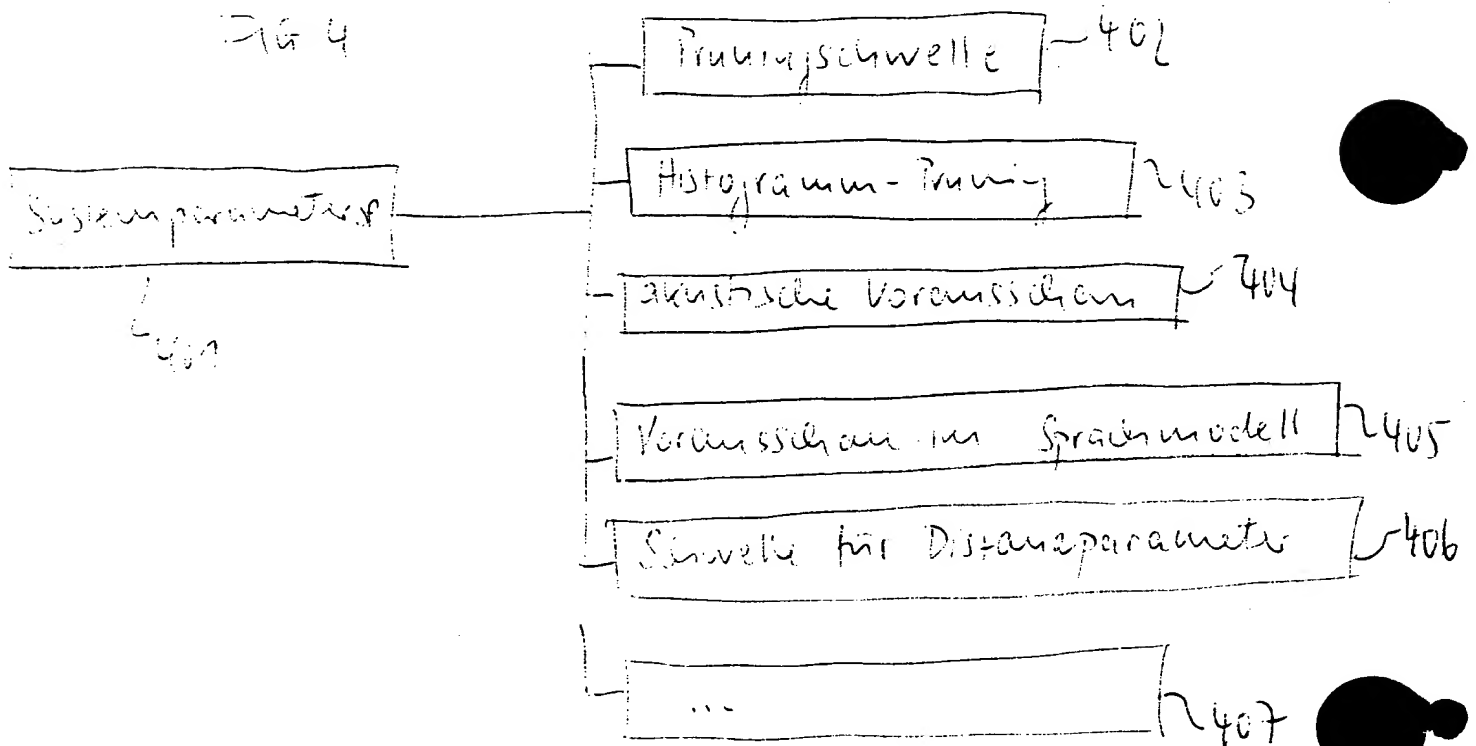




FIG 5

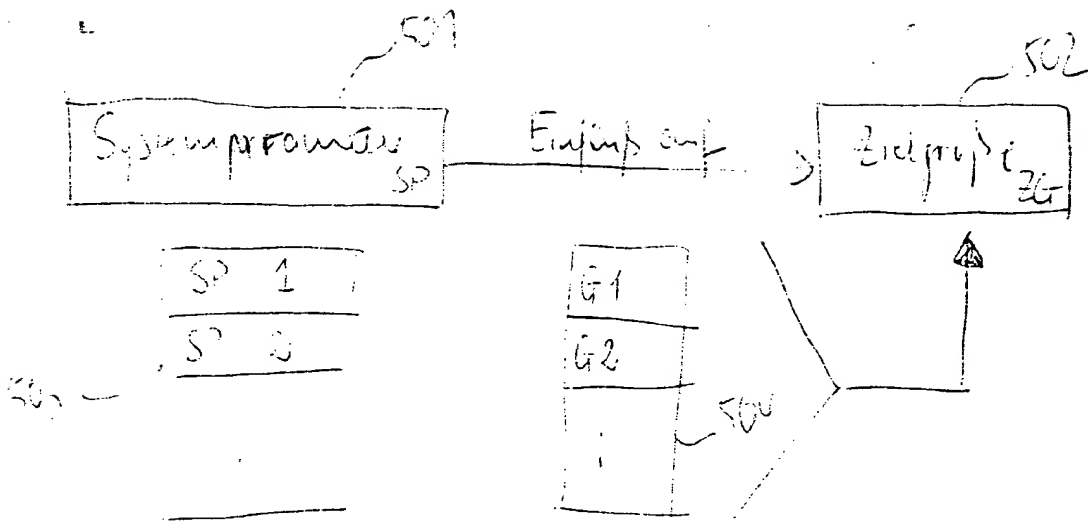


FIG 6

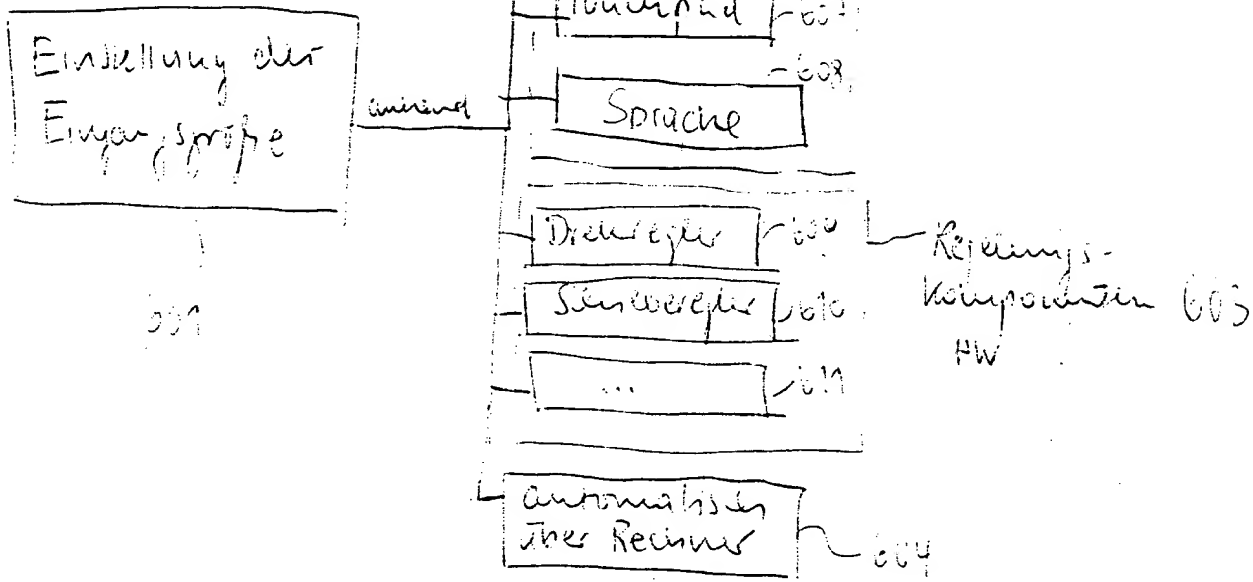
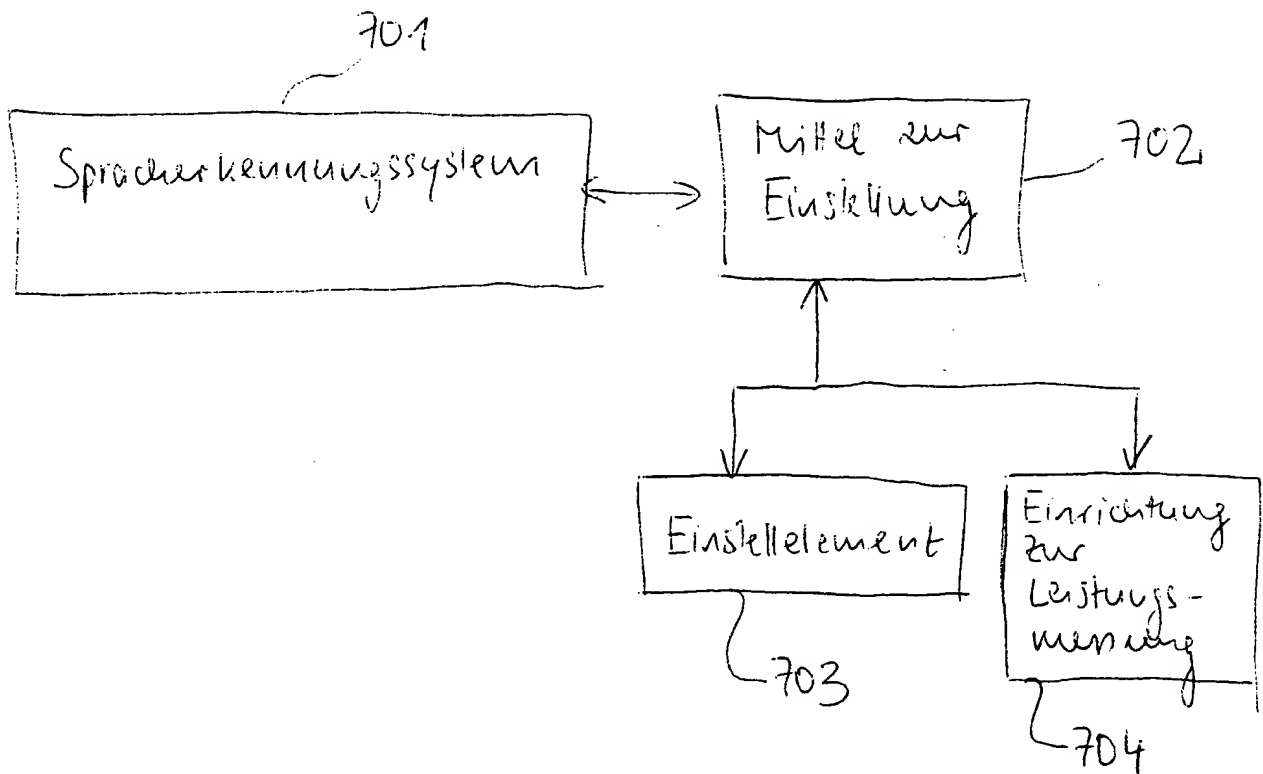


FIG 7



This Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**This Page Blank (uspto)**